DOCUMENT RESUME

ED 209 417 UD 021 831

AUTHOR Baldauf, Richard B., Jr.; Propst, Ivan K., Jr.
TITLE Assessing the Readability of Materials for Elementary ESL Pupils.
PUB DATE 24 Apr 81
NOTE 18p.; Paper prepared for the Annual Conference of the National Association for Asian and Pacific American Education (3rd, Honolulu, HI, April 24, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cloze Procedure; Difficulty Level; Elementary Education; Elementary School Students; *English (Second Language); *Objective Tests; *Readability; *Reading Comprehension; Teacher Evaluation
IDENTIFIERS *American Samoa; *Micronesia

ABSTRACT

This paper assesses the effectiveness of readability indices, standard cloze procedure, and the matching cloze procedure as determinants of the readability of supplementary English materials for elementary ESL students in a Pacific Island context. A review of readability indices and the standard cloze procedure indicated that neither procedure is an adequate measure of text readability, according to the study. Readability indices are identified as not adequately considering differences in sentence and paragraph structure, which are considered a major factor in difficulty levels of elementary ESL texts. The paper states that the standard cloze procedure is too difficult for most beginning ESL pupils because of the production skills necessary to complete the task. Matching cloze procedure, which requires primarily recognition skills, was given to elementary ESL pupils in Saipan to see whether it could be used to evaluate the readability of elementary reading materials. The matching cloze procedure is said to show excellent reliability, and generally correlated more highly with criterion measures for readability. (Author/JCD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED209417

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Richard B. Baldauf, Jr.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ASSESSING THE READABILITY OF MATERIALS FOR ELEMENTARY ESL PUPILS

By

Richard B. Baldauf, Jr.
Ivan K. Propst, Jr.

UD021831

ASSESSING THE READABILITY OF MATERIALS FOR ELEMENTARY ESL PUPILS[1]

Richard B. Baldauf, Jr.

East-West Culture Learning Institute

Ivan K. Propst, Jr.

Marianas Education Department, Saipan

ABSTRACT

A review of readability indices and the standard cloze procedure indicated that neither procedure is an adequate measure of text readability of materials suitable for elementary English as a second language (ESL) pupils. Readability indices do not adequately consider differences in sentence and paragraph structure, which are a major factor in difficulty levels of elementary ESL texts. The standard cloze procedure is too difficult for most beginning ESL pupils because of the production skills necessary to complete the task. The matching cloze procedure, which requires primarily recognition skills, was given to elementary ESL pupils in Saipan to see whether it could be used to evaluate the readability of elementary reading materials.

[1]Paper prepared for presentation at the Third National Association for Asian and Pacific American Education Conference, Honolulu, April 24, 1981.

ASSESSING THE READABILITY OF MATERIALS FOR ELEMENTARY ESL PUPILS

The purpose of this paper is to examine the readability of materials designed for use in a Pacific island context. Although English is the mother tongue for only a tiny minority of pupils in Micronesia and American Samoa, English is the medium of instruction in most Pacific island schools. Most students begin their formal English language instruction in grade one, using the Tate Oral English course. As English as a second language (ESL) pupils, they lack exposure to English, and most of the first year is devoted to developing aural-oral skills. Reading is not introduced until the second grade, or in some places even later, in these island territories.

The Oral English course, published and distributed by the South Pacific Commission (SPC), consists of 15 books which set out in detail the oral structures to be taught and the sequence and methods to be used. The SPC reading program is carefully written to parallel the structures and vocabulary learned in the oral English program. Within the program, readability is defined as those materials which parallel in form and structure the pupils' progress through the Oral English course.

While such an approach simplifies teaching and instructional decision making for those familiar with the syllabus, it has a number of limitations. First, Propst (1975) has noted that the available SPC materials are not adequate in quantity for a fully developed reading program. There needs to be a considerable amount of supplementary material provided for independent, extensive reading. Elley has further argued that although the SPC readers are attractively illustrated and contain situations familiar to Pacific island children, "they are not deliberately designed to tell a good story. There is little incentive to read just for pleasure (1980, p. 40)."

Finally, the general world of reading is not strictly controlled according to the SPC grammatical progression. Such a structured approach suggests that at least beginning materials in other school subjects, if written in English, should be developed to parallel the Oral English course. This approach was tried in the 1960's as part of the television experiment in American Samoa, but was subsequently abandoned because it was felt that it imposed language limits that were too strict for the concepts that needed to be developed in the subject areas (Baldauf 1981). Each of these limitations suggests the need to be able to establish the readability of supplementary reading materials and materials written for the content areas.

The use of existing readability formulae with elementary materials designed for use with ESL pupils is of doubtful value. Such an approach depends on counting variables such as sentence length, number of syllables, number of letters per word, regularity of phoneme-grapheme correspondence, etc., and fails to take into account conceptual difficulty, the way the text is organized, the maturity level of subjects (Lorge 1949), or the cultural and experiential background of the readers. Such formulae do not consider differences in sentence and paragraph structure nor levels of sentence embedding, which are major factors in determining the difficulty of materials for ESL pupils. Furthermore, readability formulae generate misleading normative statements of expected grade level which have been developed on and are relevant only to mainstream culture, English-speaking children.

An example to illustrate the above contention appears in a book which is indicated to have a readability grade level of 3.4. The sentence reads,

"And he was as nice as she had thought he would be." An elementary level Pacific island ESL pupil might be able to decode this sentence, but would have considerable difficulty in comprehending it. A book filled with sentences written in this manner would be generally incomprehensible.

The cloze procedure is an alternative which has been used to measure the readability of materials for ESL pupils. Anderson and Hunt (1972) used standard cloze to measure the readability of materials for ESL pupils in grades four to six in Papua New Guinea. However, our own experience with standard cloze has indicated that the procedure is too difficult for use with beginning ESL pupils (Propst and Baldauf 1979). This is because such pupils may not yet have developed the instant recall skills in English necessary to complete the task.

An alternative to the standard cloze is the matching cloze, which requires primarily recognition skills and has been shown to be an effective measure of elementary ESL reading comprehension (Baldauf and Propst 1978, Propst and Baldauf 1979). However, no work has been done with the procedure to develop a frame of reference for readability such as that available for standard cloze tests. Since matching cloze employs a procedure based on recognition rather than recall, it was expected that very different readability guidelines from those available for standard cloze would have to be developed.

Standard cloze has developed a frame of reference based on criteria of about 35 percent, 45 percent and 55 percent (Anderson and Hunt 1972, Bormuth 1967, Rankin and Culhane 1969). Materials which generate cloze test scores at 35 percent or lower are considered at the frustrational reading level; scores of about 45 percent suggest the materials are at the instructional level; while scores at 55 percent or above suggest the

materials can be read independently. Ransom (1971) has suggested similar, but lower criteria of 20/30/50 as guidelines for using cloze results with lower elementary school children.

All of these criteria have been developed by equating cloze test results to criteria first introduced by Betts (1946) of 75 percent and 90 percent correct on multiple-choice comprehension tests. Although these criteria seem to be generally accepted by reading specialists, no rationale has ever been developed for selecting these particular scores. Bormuth (1975) has developed an empirically based approach to getting criterion scores, but unfortunately his work uses standard cloze tests and native speakers of English. Therefore, the results can not be applied to the current problem.

The purpose of this study was to see if criteria could be developed for the matching cloze procedure so that the readability of materials for elementary ESL pupils could be more accurately determined.

Procedure

Matching Cloze Tests

The two matching cloze tests used in this study were based on stories taken from the South Pacific Commission's supplementary English reading materials. Each test was constructed using the six option matching cloze procedure (Baldauf and Propst 1979) and contained 50 deletions. The blanks which replaced each deletion were all of equal length. The first story, "The Red Hen and Her Eggs", was 298 words in length while the second story, "Frigate Birds", was 362 words in length. The stories were designed to be read in conjunction with books 8 and 12 respectively of the Oral English course.

Comprehension Tests

Designing multiple-choice comprehension tests for beginning ESL readers is not an easy task. Not only must the writer be aware that it is possible to write items for the same passage which vary considerably in their difficulty (Bormuth 1975, Goodman 1973), but that questions themselves may be grammatically more difficult than the passage. Furthermore, for beginning reading passages it may be difficult to find plausable item distractors due to the limited range of vocabulary being used.

In our search for a way to generate suitable multiple-choice items, we initially turned to Bormuth (1970) who had suggested a process of item transformations to overcome the item difficulty problem and to make item writing more scientific. However, our experience in trying to write items using this system confirmed Berk's (1979) assessment that cloze testing is often a preferable approach to measuring reading comprehension since item transformation procedures are very complex and are not yet well enough developed for use by classroom teachers. The discussion of the develoment of comprehension questions in other readability studies was not very helpful. While these studies usually stressed the inclusion of six or seven different comprehension skills which are assumed to define reading comprehension, they were vague in discussing the specifics of item construction.

Since it seemed to be impractical to construct psychometrically sound multiple-choice comprehension questions, it was decided to write 30 yes/no (i.e., true-false) items for each of the comprehension passages. The pupils were familiar with this format from classroom exercises. This type of item also reduces to a minimum the problem of items (i.e., questions) being grammatically more difficult than the reading passage. Half the items were written with positive responses and half with negative responses.

Where possible, items for both tests were developed using sentence paraphrases. Anderson (1972) argues convincingly that a student may be able to answer rote items derived from a base sentence transformation without comprehending the statement. Items using sentence paraphrases, provided the changes do not make the new sentence more difficult than the original, insure that comprehension is being measured.

Perceived Readability

Since some doubts exist about the suitability of traditional comprehension tests as criterion for judging the readability of materials for beginning ESL readers, we decided to explore some group oriented, alternative ways to estimate readability. More individually oriented methods such as mis-cue analysis or story retelling were felt to be inappropriate for use in an initial exploratory study such as this one.

Teachers' perceptions of readability were obtained by getting them to rank order all the pupils in their class from the best reader to the poorest one. Then they were shown an unmutilated version of the cloze passage their students were going to read. After reading the passage, the teachers were asked to divide the rank ordering of students they had previously completed into three groups. The groups were to contain pupils for whom the passage would be (1) suitable as an extensive (independent) piece of reading, (2) suitable as an intensive (instructional) piece of reading, and (3) not suitable or too difficult (frustrational) for these pupils to read and comprehend.

Pupil perceptions about the readability of the passages were collected by asking each pupil to choose one of the following statements about the story (s)he had just read:

(1) Stories like this are quite easy. I can read stories like this by myself.

(2) Stories like this are a bit hard. I would like to read stories like this in class with a friend or my teacher; or

(3) Stories like this are very hard. I would like to read an easier story than this one.

To be sure the pupils understood what they were being asked to do, the instructions and the three responses were explained in both English and Chamorro.

Subjects

One hundred and fifty-four ESL pupils initially participated in this study, 79 from grade four, and 75 from grade five. They were the complete classes in those grade levels in an academically average elementary school located in the major population center on the island of Saipan in the Northern Mariana Islands. Due to incomplete data on one or more of the four variables, only 56 fourth graders and 63 fifth graders were included in the final analysis of data.

Administration

The matching cloze exercise was administered first to prevent the pupils from reading the unmutulated versions of the paragraphs before attempting to fill in the missing words. Pupils in grade four took "The Red Hen and Her Eggs" while pupils in grade five took the "Frigate Birds". One week later the comprehension exercise was administered. Both exercises were given by the classroom teachers to de-emphasize the testing situation. Both exercises were untimed and students had ample time to complete them. Since Bormuth (1967) has shown that order of administration does not significantly alter test results, a counter-balanced test administration design was not used.

Teacher ratings of pupils' reading comprehension and their estimates of passage readability were collected prior to the start of testing. Student estimates of passage difficulty were collected upon the completion of both reading tasks.

Results

Means, standard deviations, and KR_{20} reliability estimates for the cloze and comprehension tests are set out in Table 1. The table also includes the proportion of pupils who were rated as independent readers by teachers and by their own self ratings. The correlation between each of the four variables is also included in the table.

INSERT TABLE 1 ABOUT HERE

Comparable scores on matching cloze and the comprehension test for grade 4 were 92.85 for Betts' criteria of 90% and 67.76 for her criteria of 75%. The standard error was 4.00. Only the comprehension test results were used because matching cloze correlated highly (r=.874) with that result, but only .261 and .467 with teacher and student ratings of independent reading proficiency respectively.

Fifth grade scores on matching cloze and the comprehension test were judged not to be comparable because of low the reliability (r=.65) for the latter measure and its low correlation (r=.633) with matching cloze. Using the combined criteria of the comprehension test and teacher ratings a somewhat improved multiple correlation of .715 was used to predict a matching cloze score of 70.34 at the 75% criterion level. However, the predicted score of 105.67 (an impossible result) for the 90% criterion reflects the high level of error still found in this model. A

better estimate of the 90% criterion may be the 88.33 mean score for 5th grade pupils rated by teachers at an independent reading level.

Discussion and Conclusions

The problems encountered in developing readability criteria for matching cloze procedure used with elementary ESL pupils, reflect the difficulty finding accurate measures of readability for these pupils. The matching cloze procedure showed excellent reliability and generally correlated more highly with the other three measures (i.e., was more valid) than those measures did among themselves. However, the other measures (i.e., the criteron) were not as satisfactory.

The two comprehension tests were limited by their true-false format which includes a large guessing component (i.e., Scores less than 15 out of 30 are unlikely if students answers all the items). This problem was compounded by the lay out of the tests which required pupils to read relatively long sections of about 100 words and then answer ten items. This procedure probably required too much student recall. The tests could have been improved by reducing the length of the passage and by reducing the number of items in each section to perhaps 60 words and 6 items. Shorter passages might have encouraged pupils to go back and restudy the text. This should reduce guessing and perhaps improve test results. This problem was especially evident for 5th graders where only two pupils had scores higher than 26/30.

The accuracy of teacher ratings of pupils into independent, instructional, and frustrational readers was hampered by a tendency to over-estimate pupils reading ability and by a reluctance to use the third category frustrational reader. There was also a natural tendency, despite instructions to the contrary, to place some students in each of the three categories within a

classroom. This was particularly inappropriate as the classes are homogeneously grouped in that school.

The accuracy of pupil ratings of themselves was effected by student pride. Many pupils were reluctant, despite assurances that this was not a "test", to admit they couldn't read the passages. However, while the both sets of rating correlations are low, many of the individual ratings agree closely with the test results. The overall correlations are reduced by a few students or teachers ratings which are completely inappropriate as judged by the rest of the data.

Despite the difficulties encountered in collecting completely satisfactory data, the results do provide some tentative guidelines about readability of ESL materials as predicted by the six option matching cloze procedure. Pupils with scores of about 90% or better on such tests can probably read such material independently. Scores of between 70% and 90% suggest instructional use of the materials while scores below 70% probably indicate frustrational readers. These results should now be validated with individual pupils using miscue analysis or story retelling procedures.

Despite the difficulties encountered, we believe the use of matching cloze readability criteria is worth pursuing. Score guidelines of 65/80/95 percent should be more acceptable to both students and teachers than the standard cloze results of 35/45/55 or 20/30/50 cited earlier in this paper.

TABLE 1

Means, Standard Deviations, KR_{20} Reliabilities, and Correlations by Grade for Four Measures of Text Readability

Sample Measure	Mean	S.D.	KR_{20}	Correlation Matrix[1] M-C	Comp	Tch	Pupil
Grade 4 (N=56)							
Matching Cloze	70.64	28.32	.98	1.000	0.874	0.261	0.467
Comprehension	23.02	5.81	.87		1.000	0.164	0.467
Teacher-Indep.	.45	NA	NA			1.000	0.301
Pupil-Independ.	.55	NA	NA				1.000
Grade 5 (N=63)							
Matching Cloze	66.16	28.54	.98	1.000	0.633	0.610	0.293
Comprehension	17.87	4.41	.65		1.000	0.520	0.293
Teacher Indep.	.38	NA	NA			1.000	0.165
Pupil Independ.	.40	NA	NA				1.000

[1]The matching cloze - comprehension correlations are product-moment correlations, the teacher/pupil - matching cloze/comprehension correlations are point biserials, while the teacher-pupil results are phi-coefficients.

REFERENCES

Anderson, Jonathan and A. H. Hunt. "A Frame of Reference for Cloze Tests of Readability of English learned as a Foreign Language." Papua New Guinea Journal of Education, vol. , no. (1972), pp. 184-188.

Anderson, R. C. "How to Construct Achievement Tests to Assess Comprehension." Review of Educational Research, vol. 42, no. 2 (Spring 1972), pp. 145-170.

Baldauf, Richard B., Jr. "Educational Television, Enculturation, and Acculturation: A Study of Change in American Samoa." International Review of Education, vol. 27, no. 3 (1981)

Baldauf, Richard B., Jr. and Ivan K. Propst, Jr. "Preliminary Evidence Regarding the Validity of a Modified Cloze Procedure for Lower Elementary ESL Students." Educational and Psychological Measurement, vol. 38, no. 2 (Summer 1978), pp. 451-455.

Baldauf, Richard B., Jr. and Ivan K. Propst, Jr. "Matching and Multiple-Choice Cloze Tests." Journal of Educational Research, vol. 72, no. 6 (July-August 1979), pp. 321-326.

Berk, Ronald. "The Relative Merits of Item Transformations and the Cloze Procedure for the Measurement of Reading Comprehension." Journal of Reading Behavior, vol. 11, no. 2 (Summer 1979), pp. 129-138.

Betts, E. A. Foundations of Reading. New York: American Book Co., 1946.

Bormuth, John R. "Comparable Cloze and Multiple-Choice Comprehension Test Scores." Journal of Reading, vol. 10, no. 5 (February 1967), pp. 291-299.

Bormuth, John R. On the Theory of Achievement Test Items. Chicago: University of Chicago Press, 1970.

Bormuth, John R. "The Cloze Procedure: Literacy in the Classroom." Help for the Reading Teacher: New Directions in Research. W. D. Page Ed., pp. 60-90. Urbana Ill.: Eric Clearinghouse on Reading, 1975.

Elley, Warwick B. "A Comparison of 'Content-Interest' and 'Structuralist' Reading Programmes in Niue Primary Schools." New Zealand Journal of Educational Studies, vol. 15, no. 1 (May 1980), pp. 39-53.

Goodman, Kenneth S. "Testing in Reading: A General Critique." Accountability and Reading Instruction. Critical Issues. Robert B. Ruddell Ed., pp. 21-23. Urbana, Ill.: National Council of Teachers of English, 1973.

Lorge, Irving. "Readability Formulae - An Evaluation." Elementary English, vol. 26, no. 2 (February 1949), pp. 86-95.

Rankin, Earl F. and Joseph W. Culhane. "Comparable Cloze and Multiple-Choice Test Scores." Journal of Reading, vol. 13, no. 3 (December 1969), pp. 291-299.

Propst, Ivan K., Jr. "A Serious Deficiency in the SPC English Language Programme." South Pacific Bulletin, vol. 25, no. 2 (2nd quarter 1975), pp. 24-27.

Propst, Ivan K., Jr. and Richard B. Baldauf, Jr. "Use Matching Cloze Tests for Elementary ESL Pupils." The Reading Teacher, vol. 32, no. 6 (March 1979), pp. 683-690.

Ransom, Peggy E. "Determining Reading Levels of Elementary School Children by Cloze Testing." Forging Ahead in Reading. J. A. Figurel Ed., pp. 477-482. IRA 12th Annual Convention Procedings, vol. 21, pt. 1 (1971).

READING COMPREHENSION

Exercise 2A

Name _____ Date _____

School _____ Grade _____

EXAMPLE

Going to School

Tom ___is___ going to school.

Mary is _____ to school, too.

She's _____ to Tom.

Mary: "Where's Anna, Tom?"

Tom: "Look. _____ there.

She's running to _____."

going
is
school
she's
talking
to

Now Peter is walking with Tom and Mary is walking with Anna.

_____ are they going?

_____ going to school.

They're _____ fast.

Peter and Tom _____ clean.

Mary and Anna are _____, too.

are
clean
they're
walking
where
with

Exercise 2B

Name _____ Date _____

School _____ Grade _____

EXAMPLE

Going to School

Tom is going to school.
Mary is going to school, too.
She's talking to Tom.
Mary: "Where's Anna, Tom?"
Tom: "Look. She's there.
She's running to school."
Not Peter is walking with Tom and
Mary is walking with Anna.
Where are they going?
They're going to school.
They're walking fast.
Peter and Tom are clean.
Mary and Anna are clean, too.

Circle the correct answer.

1. Tom is walking to school. Yes No
2. Mary is talking to Tom. Yes No
3. Peter is walking with Anna. Yes No
4. The children are walking fast. Yes No
5. Tom is running to school. Yes No
6. The children are dirty. Yes No